AD_____

Award Number: DAMD17-02-1-0398

TITLE: Discovery of Novel Metastasis Genes in Breast Cancer

PRINCIPAL INVESTIGATOR: Anton Wellstein, M.D., Ph.D.

CONTRACTING ORGANIZATION: Georgetown University
Washington, DC 20007

REPORT DATE: July 2004

TYPE OF REPORT: Annual

PREPARED FOR: U.S. Army Medical Research and Materiel Command
Fort Detrick, Maryland 21702-5012

DISTRIBUTION STATEMENT: Approved for Public Release;
Distribution Unlimited

The views, opinions and/or findings contained in this report are those of the author(s) and should not be construed as an official Department of the Army position, policy or decision unless so designated by other documentation.

REPORT DOCUMENTATION PAGE

Form Approved
OMB No. 074-0188

1. AGENCY USE ONLY (Leave blank)	2. REPORT DATE	3. REPORT TYPE AND DATES COVERED
	July 2004	Annual (1 Jul 03 - 30 Jun 04)

4. TITLE AND SUBTITLE
Discovery of Novel Metastasis Genes in Breast Cancer

5. FUNDING NUMBERS
DAMD17-02-1-0398

6. AUTHOR(S)
Anton Wellstein, M.D., Ph.D.

7. PERFORMING ORGANIZATION NAME(S) AND ADDRESS(ES)
Georgetown University
Washington, DC 20007

E-Mail: wellstea@georgetown.edu

8. PERFORMING ORGANIZATION REPORT NUMBER

9. SPONSORING / MONITORING AGENCY NAME(S) AND ADDRESS(ES)
U.S. Army Medical Research and Materiel Command
Fort Detrick, Maryland 21702-5012

10. SPONSORING / MONITORING AGENCY REPORT NUMBER

11. SUPPLEMENTARY NOTES

12a. DISTRIBUTION / AVAILABILITY STATEMENT
Approved for Public Release; Distribution Unlimited

12b. DISTRIBUTION CODE

13. ABSTRACT (Maximum 200 Words)

Under this proposal we plan to identify genes that drive the homing of invasive primary breast cancer to its metastatic sites.

For the discovery of such genes, in vivo phage display of breast cancer cDNA libraries is used. Validation is achieved by expression monitoring in histological sections as well as by functional assays. In the current series of studies we validated the breast cancer library. We obtained the first results from the in vivo selection. We report those data here

14. SUBJECT TERMS
Breast Cancer

15. NUMBER OF PAGES
52

16. PRICE CODE

17. SECURITY CLASSIFICATION OF REPORT	18. SECURITY CLASSIFICATION OF THIS PAGE	19. SECURITY CLASSIFICATION OF ABSTRACT	20. LIMITATION OF ABSTRACT
Unclassified	Unclassified	Unclassified	Unlimited

Table of Contents

INTRODUCTION

The major problem of clinical cancer is metastasis to distant organs [1-3]. It has been known for over 100 years that there is a close relationship between "seed and soil" in terms of organ-specific metastasis and we now have the tools available to potentially identify tumor cell-derived proteins that aid in the organ-specific metastasis. Here we plan to isolate from human breast cancer cells genes that are candidates for the homing in of tumor cells to blood vessels in distinct organs of metastasis. This will be accomplished using phage display of cDNA libraries and will identify protein fragments that serve as ligands for the local homing in.

In addition to our better understanding of tumor cell metastasis in the body, discoveries made under this proposal will allow us to utilize such genes for both diagnostics and prognostic as well as a therapeutic targeting in the future.

Task 1: Identification of protein fragments that allow for organ-specific metastasis.

Task 2: Study expression patterns of candidate homing genes.

Task 3: Study the functional role of candidate homing genes.

BODY

Work accomplished during the current award cycle:

This work is related to Task 1 and Task 2:

During the current funding cycle, we selected for bone marrow, brain and lung homing genes using a phage library generated from MDA-MB231 breast cancer cells. For this, the library was titrated extensively, tested for insert size gloablly and individual plaques were analyzed. Then we continued with the selection in mice.

Quality control experiments:

For this, we first assured that we can rescue the phage from tissue homogenates and we used discarded tissues from other experiments to assess that.

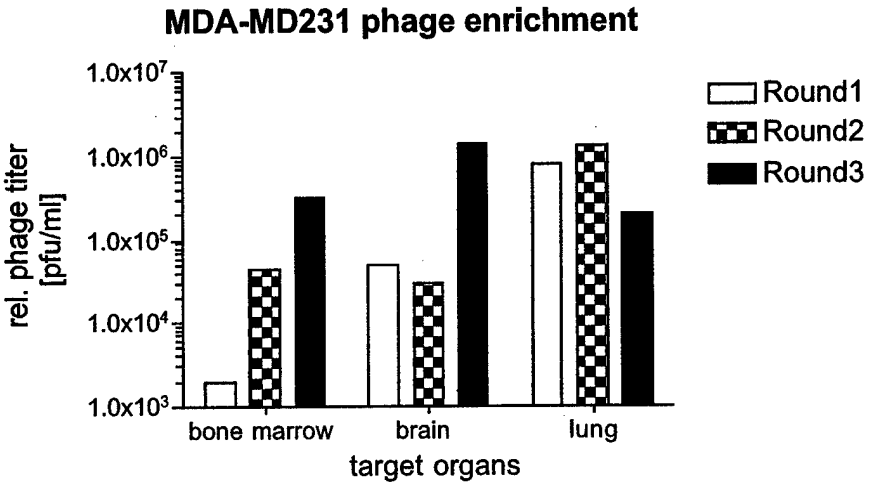
In vivo selection for organ selective homing phage:

We then then went forward with the in vivo selection by injection into the tail vein and short term circulation of phage in control mice that were to be sacrificed due to the termination of the respective studies.

Figure 1 shows to what extent we have then been able to enrich for phage particles after injection of a MDA-MB231 breast cancer cell phage display library into a series of mice for each of the particular target organ selections. One mouse per round was used for selection of homing phage to specific target organs. A series of organs were then harvested and titered for retained phage. Absolute titers are given for each target organ and we wish to point to the log scale ! The most striking is the result with the bone marrow and the brain, where we saw a 100-fold enrichtment. The selection towards the lungs appears to have been positive after the first round and not increase over the next ones.

Figure 2 gives a quality of the phage rescued after round 3. For this the inserts were amplified by PCR using primers hybridizing to the phage flanking arms. Obviously, distinct bands can be identified suggesting enrichment of a limited number of distinct genes.

Figure 3 shows the picking of individual plaques and the amplification of the inserts from the phage harvested from the different target organs. Interpretation: We appear to have selected a set of genes that are in part repeated based on the insert sizes.



target organs	Round1	Round2	Round3
bone marrow	2.00e+03	4.40e+04	3.26e+05
brain	5.00e+04	3.00e+04	1.40e+06
lung	7.88e+05	1.30e+06	2.04e+05

Figure 1: Phage particles retained after different rounds targeting different organs, i.e. bone marrow, brain and lungs with a phage library generated from MDA-MB231 breast cancer cells. 10^8 phage were injected. Absolute number of phage rescued are given for the respective organ after each round of in vivo selection.

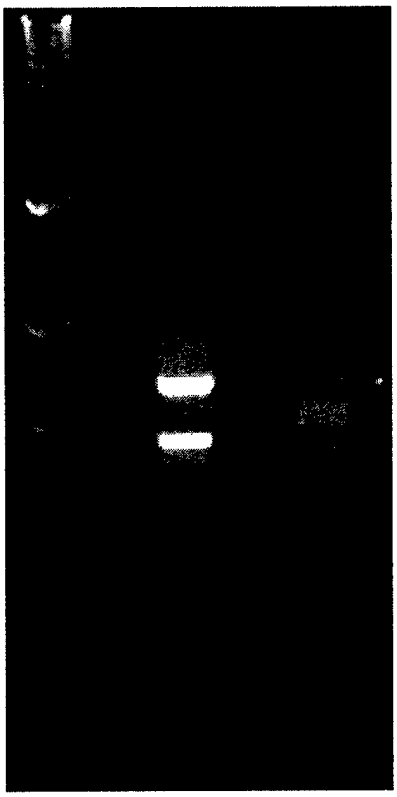


Figure 2. Amplification of phage inserts using phage rescued from different organ systems as the template. Left to right: Lane 1 = Markers. Lane 2 = bone marrow, Lane 3 = brain, Lane 4 = lung. The top insert harvested from the bone marrow is approximately 750 bp, the two strong inserts from the lungs are between 400 and 600 bp and the range of bands from the lungs is 300 to 550 bp.

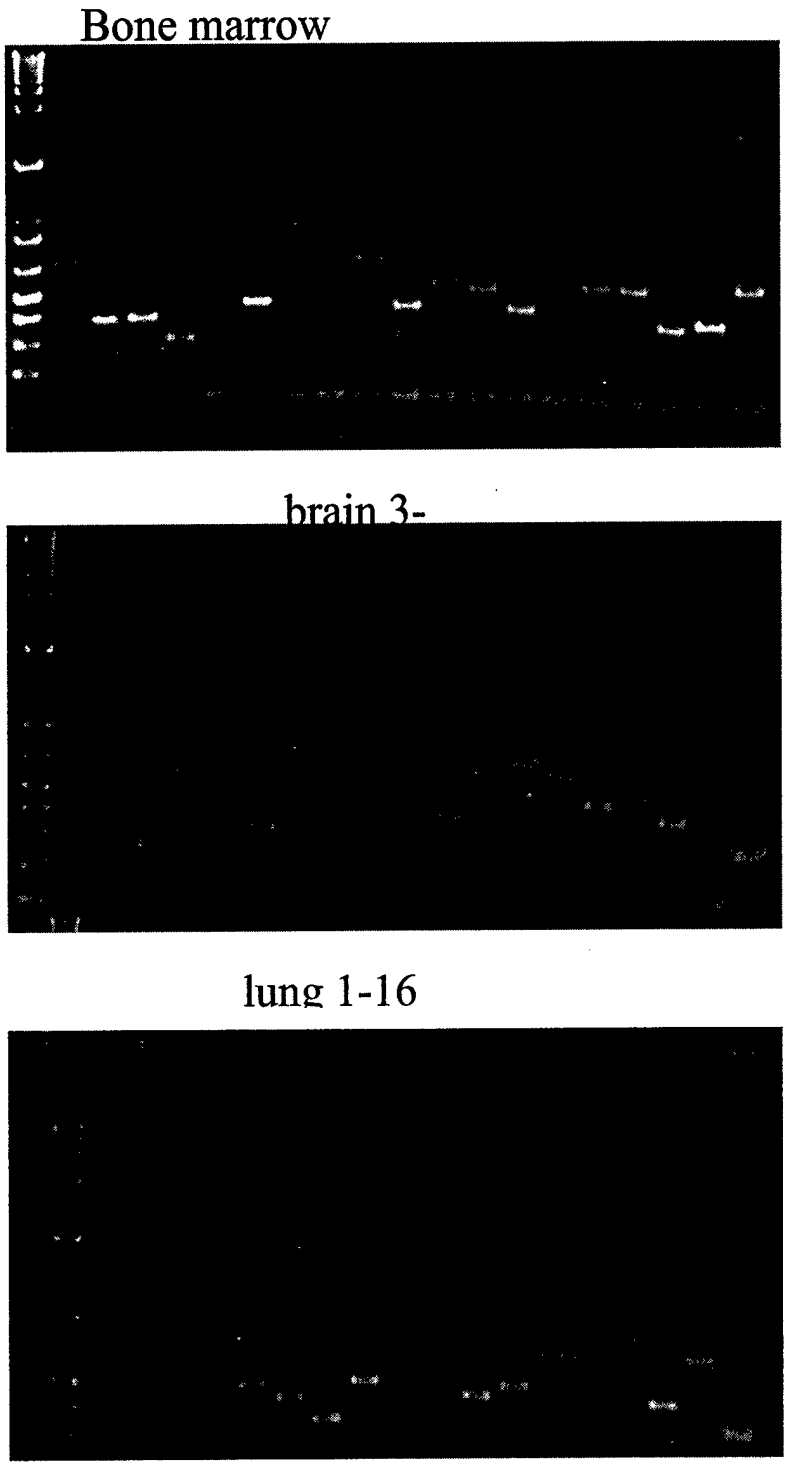


Figure 3: Selection of individual plaques selected from phage harvested from the respective target organ and amplification of the respective inserts.

Methods:

The phage library was generated in T7 phage by insertion of size selected cDNA from MDA-MB231 cells (>200 bp) into the 10B gene of T7.

This library was amplified and particles at 10^8 titer were injected into the tail vein of fully anesthetized mice, and allowed to circulate briefly. Target organs were harvested after draining the blood from the target organ and washing out residual blood with saline perfusion through the left ventricle.

Organs were then homogenized and phage recovered by addition of E. Coli. Phage particle number was titrated by plating on agar plates.

Expanded phage were re-injected for subsequent rounds.

Phage inserts were amplified using primers hybridizing to the flanking regions of the 10B gene. The products were run on agarose gels and visualized by DNA stain.

INTERPRETATION OF THE DATA.

We conclude that have selected a population of phage that drives bone marrow and brain selective homing. We are not as sure about the lung results yet.

KEY RESEARCH ACCOMPLISHMENTS

1. We established selection of organ selective homing of phage from a human breast cancer cell line and can show organ selective enrichment for the bone marrow and the lungs.

REPORTABLE OUTCOMES

Henke et al. Attached manuscript accepted for publication with minor revisions.

CONCLUSIONS

- Organ selective homing genes can be selected by in vivo selection of phage displaying genes from breast cancer cells.
- Sequence analysis of the genes will show, what their known function is as to what extent that may contribute to their role in organ selective homing.

REFERENCES

1. Zetter, B. Endothelial cell heterogeneity: Influence of vessel size, organ localization, and species specificity on the properties of cultured endothelial cells. in *Endothelial Cells* (ed. Ryan, U.) 63-79 (CRC Press, New York, 1988).
2. Blood, C.H. & Zetter, B.R. Tumor interaction with the vasculature: Angiogenesis and metastasis. *Biochim.Biophys.Acta* **1032**, 89-118 (1990).
3. Liotta, L.A., Steeg, P.S. & Stetler-Stevenson, W.G. Cancer metastasis and angiogenesis: An imbalance of positive and negative regulation. *Cell* **64**, 327-336 (1991).

APPENDICES

Henke et al. Manuscript accepted for publication

Gene expression analysis in sections and tissue microarrays of archival tissues by mRNA *in situ* hybridization.

Ralf T. Henke[1], Anirban Maitra[2], Soonmyung Paik[3] and Anton Wellstein[1]

[1] Lombardi Cancer Center, Georgetown University, Washington, DC, USA

[2] Department of Pathology, Johns Hopkins University School of Medicine, Baltimore, MD, USA

[3] Division of Pathology, National Surgical Adjuvant Breast and Bowel Project (NSABP), Pittsburgh, PA, USA

Correspondence to:

Anton Wellstein, MD/PhD

Georgetown University, Lombardi Cancer Center, Research Building E311

3970 Reservoir Rd. NW, Washington, DC 20057, USA

Phone: USA (202) 687-3672

Fax: USA (202) 687-4821

Email: wellstea@georgetown.edu

Running Title: mRNA ISH of archival tissues

Keywords: mRNA, *in situ* hybridization, tissue microarry, expression analysis

Abbreviations: AIB1, amplified in breast cancer 1; FGF-BP, fibroblast growth factor binding protein; ALK, Anaplastic lymphoma kinase; PTN, pleiotrophin; ISH, in situ hybridization; IHC, immunohistochemistry; TMA, tissue microarray;

Summary

Altered expression of genes in diseased tissues can prognosticate a distinct natural progression of the disease as well as predict sensitivity or resistance to particular therapies. Archival tissues from patients with a known medical history and treatments are an invaluable resource to validate the utility of candidate genes for prognosis and prediction of therapy outcomes. However, stored tissues with associated longterm follow-up information typically are formalin-fixed, paraffin-embedded specimen and this can severely restrict the methods applicable for gene expression analysis. We report here on the utility of tissue microarrays (TMAs) that use valuable tissues sparingly and provide a platform for simultaneous analysis of gene expression in several hundred samples. In particular, we describe a stable method applicable to mRNA expression screening in such archival tissues. TMAs are constructed from sections of small drill cores, biopsied from tissue blocks of archival tissues and multiple samples can thus be arranged on a single microscope slide. We used mRNA *in situ* hybridization (ISH) on >500 full sections and >100 TMAs for >10 different cDNAs that yielded >10,000 data points. We provide detailed experimental protocols that can be implemented without major hurdles in a molecular pathology laboratory and discuss quantitative analysis and the advantages and limitations of ISH. We conclude that gene expression analysis in archival tissues by ISH is reliable and particularly useful when no protein detection methods are available for a candidate gene.

Introduction

The number of genes or cDNAs that are being identified as potential contributors to different diseases is rapidly increasing due to ever expanding gene expression screening with cDNA microarrays. Expression analyses of chosen candidate genes in a large series of tissue samples are then of crucial importance to address the relevance of such candidates as well as translation to routine applications once that relevance is established. In general, expression of a gene can be investigated at the mRNA or protein level. However, protein expression analysis is frequently limited by the lack of a suitable antibody, leaving only the option of mRNA studies. This problem is compounded if only cDNA fragments are known (e.g. ESTs) or have been identified.

Even if protein analyses can be performed, mRNA studies still provide important additional information, for example by detailing changes in the efficiency of protein translation during different steps of malignant progression of a tumor or with regard to different developmental or hormonal stages (Steel et al., 1988). The analysis of the mRNA in conjunction with protein expression also allows further quality control for both methods. mRNA studies may also serve to identify the cells of origin for secreted proteins.

Although initial studies can often be performed with fresh frozen material, the bulk of any archive of a pathology institute or a clinical study archive, like that of the NSABP, is comprised of tissue samples fixed with formalin and embedded in paraffin. Therefore, the most informative and precious samples (for which clinical follow-up and diagnostic, demographic, therapeutic etc. information are available) will often be only available as paraffin blocks. In order to achieve maximum efficiency and stability, any large-scale expression analysis on tissues should also consider the use of tissue microarrays (TMA), constructed from paraffin embedded "donor" tissue blocks to avoid depletion of resources. To approach this we will address several crucial points in this papers:

1. Utility of formalin-fixed, paraffin-embedded archival tissues.
2. The ability to distinguish expression in different cell types in heterogenous tissues.
3. Sensitivity and specificity.
4. Applicability of the same protocol for different genes in different different tissues.
5. Application to tissue microarrays (TMAs) for high throughput studies.
6. Evaluation of results from TMAs

1. Formalin-fixed, paraffin-embedded archival samples for RNA studies

When using tissue for the ISH or for other histological application, it is necessary to consider the processing of the tissue prior to the actual experiment, in particular, fixation and embedding. Common fixation methods include freezing (liquid nitrogen), ethanol and a variety of formalin-based solutions. Unless tissue is frozen, a durable embedding must be performed. Paraffin will be used most frequently but special plastic formulations are an alternative.

When investigating mRNA, fresh frozen tissue is considered by many investigators as the best choice. However, such tissues are limited and rarely have extensive pathological or long-term clinical follow-up data linked to them. In contrast, formalin-fixed, paraffin-embedded archival tissue typically will have more long-term information associated. Studies to assess the relevance of a gene or mRNA will therefore, at some point, face the necessity to investigate sections from such archival sources. Whether mRNA in such tissues is preserved sufficiently and with good representation is of concern although this also applies to proteins. In general, formalin-fixed paraffin-embedded samples have been shown to be useful for RNA studies and to produce stable results compared to other fixation techniques (Bismar et al., 2003; Fink et al., 1998; Jin et al., 1999; Kabbarah et al., 2003; Van Deerlin et al., 2002). Typically, mRNA fragments of up to 100 bases can be quantified from paraffin tissues by RT-PCR with good sensitivity and specificity. Longer consecutive sequences prove more difficult to detect. This is likely due to mRNA fragmentation during the extraction

process and may not apply as much to *in situ* hybridization which leaves the targeted mRNA in its locus, i.e. "in situ". We usually design relatively large probes (>500 bases) for *in situ* hybridization to achieve maximum hybridization efficiency with variable mRNA fragments and thus try to minimize the impact of mRNA fragmentation. Our experience with a series of probes in very different tissues is presented below.

2. Methods to investigate RNA expression

Common ways to analyze mRNA expression in tissue samples include Northern blot, reverse transcription with consecutive PCR (RT-PCR) or quantitative real time RT-PCR, *in situ* RT-PCR and *in situ* hybridization (ISH) for mRNA as well as cDNA microarrays. Major concerns are sensitivity and specificity, tissue heterogeneity of samples (i.e. tumors intermixed with normal stroma) and heterogeneous expression within one cell type in different areas of a sample. Unless laser capture micro dissection is performed (Fend and Raffeld, 2000) to isolate a particular cell type (for example cancer cells), or an *in situ* method is used, contamination with other cell types often leads to problems in interpreting the results in a "pooled RNA" assay by the other methods.

In situ methods allow the evaluation of mRNA expression separately for particular cell types and tissue compartments due to the preservation of the histological architecture. For example, we observed –quite unexpectedly- that the receptor for the growth factor pleiotrophin (PTN), Anaplastic Lymphoma Kinase (ALK) is expressed in neovascular endothelial cells in pancreatic adenocarcinoma, glioblastoma and colon carcinoma (unpublished data). Endothelial cells of normal vessels did lack expression. This information would have been lost when using a Northern blot or RT-PCR from tissue homogenates and the expression would have been attributed to the tumor cells, which also express the ALK receptor.

Due to the preservation of the histology, the ISH also allows for relatively easy identification of artifacts, low tissue quality, and necrotic and autolytic tissues. In addition, the mRNA expression

in cancer cells can be compared selectively with that of the appropriate reference cells rather than a "whole organ background". For example, when investigating breast cancers or pancreatic adenocarcinoma, epithelial cells in breast tissue and those of the small pancreatic ducts, respectively, can be identified and used as a reference. Stroma cells and – in the case of pancreas – acinar cells, which are not the cells of origin for the mentioned cancers, can be easily excluded from the analysis as well as those samples, which did not contain appropriate reference cells.

Inspite of these advantages of the *in situ* hybridization method, a some issues need serious consideration: Although a semi-quantitative evaluation of expression levels can be achieved by assessing staining intensity and frequency, real time PCR will allow for quantitation of expression over a wider range with more precision. Also, protein studies will obviously generate qualitatively distinct information location, post-translational modification and possibly activation status.

3 and 4. Overview of the *in situ* hybridization

Recent work shows that ISH has advantages in sensitivity and specificity over *in situ* RT-PCR (Steel et al., 2001) further suggesting that the ISH is an established and reliable method. Two types of probes are established for the mRNA *in situ* hybridization of frozen or paraffin embedded tissue: Radioisotope-labeled (^{33}P or ^{35}S) and non-isotopic (e.g. digoxigenin labeled) probes. Digoxigenin-labeling methods do provide a higher resolution of the signal localization with comparable sensitivity and specificity (Steel et al., 1998). In addition, the problems of handling and disposal of radioactive isotopes are eliminated. Also, with the right chromogen to visualize Digoxigenin labeled probes, signals can still be evaluated after a long time, whereas the half-life of isotopes is limited (24 days for ^{33}P, 87 days for ^{35}S) and requires higher experimental effort to preserve the signal (for example by photographic film). We used Digoxigenin labeled probes for recent studies and found them to generate better quality results and easier to handle in comparison to

radioactive probes (e.g. for FGF-BP: Digoxigenin (Kagan et al., 2003; Ray et al., 2003) versus radioactive (Kurtz et al., 1997).)

Detection of the Digoxigenin after the probe hybridization is usually performed with antibodies or antibody-fragments. These can either be labeled directly with an enzyme (HRP or AP), biotin or flourescein/rhodamine. Alternatively, unlabelled primary anti-Digoxigenin antibodies can be detected with an appropriate secondary antibody.

Flourescein based systems have the lowest sensitivity and should be used only for high abundance mRNAs. For most mRNAs, we observed good sensitivity with AP conjugated FAB-fragments. Low abundance mRNAs, which cannot be detected with AP, can sometimes be visualized by HRP labeled FAB-fragments with consecutive tyramide signal amplification. However, this protocol is significantly more expensive and usually not required if the hybridization is first optimized by using long mRNA probes at high concentration.

The ISH method reported here is based on a protocol reported by Panoskaltsis-Mortari and Bucy (1995) and has been successfully used in our lab for >500 slides, including more than 100 tissue microarrays and more than ten different mRNAs/cDNAs have been studied. An overview of the genes and tissues studied so far is given in Table 1. Some of the data are already published (Kagan et al., 2003; Klomp et al., 2002; Kurtz et al., 1997; List et al., 2001; Liu et al., 2001; Mashour et al., 1999; Powers et al., 2002; Schulte et al., 2000) or are submitted/under revies (Henke et al.).

5. Tissue microarray overview

The use of tissue microarrays (TMAs) (Kononen et al., 1998) for ISH adds further power to any histological screening approach and thus makes an ISH even more attractive, compared to Northern blot and PCR (Bubendorf et al., 2001). This is particularly true, when investigating a novel

cDNA of interest. In brief, a TMA contains a large number of cylindrical tissue cores from paraffin embedded full sections ("donor" blocks) arrayed into a "receptor" paraffin block.

The first obvious advantage of TMAs is that a large number of samples can be investigated with less experimental effort while retaining the previously discussed features of the ISH. In addition, TMA allow for optimal method control since all samples (cores) on a TMA slide will have the exact same conditions before and during the experiment. In particular, all samples (cores) will have had the same time between sectioning and staining, the same hybridization conditions and the same concentrations of probe, antibodies, enzyme and chromate (Bubendorf et al., 2001). Therefore, the use of a TMA adds further advantages with regard to the evaluation of the observable maximum and minimum expression levels, thus eliminating a slide-to-slide variability often seen when investigating a series of regular sections. Thus, categorization of expression levels of all cores can be easily performed. In addition, cores with artifacts and low tissues quality or mRNA degradation can be quickly identified by comparison with other cores since other reasons for the observed variations due to the method are eliminated. Furthermore, since up to 200 sections can be cut from a TMA block (Kononen et al., 1998), comparison of the expression of different mRNAs or proteins on parallel sections can be achieved with considerably less effort for a large number of genes or cDNAs. We therefore began using TMAs very early for ISH in order to maximize the number of samples and the stability of results.

6. Evaluation of ISH staining results for TMAs

Obviously, it is of concern whether the small tissue sample present in a TMA is representative for the original full section. This concern has been addressed in different studies showing that results from a TMA with three of four sample cores per tissue are highly representative compared to results of full sections (Camp et al., 2000; Nocito et al., 2001; Simon et al., 2004; Simon et al., 2001). For some investigations, as few as two cores were sufficient (Camp et al., 2000)

though we strongly recommend a redundance of four cores from our experience, due to the frequent loss of cores or poor quality of a given core. In addition, it is worth mentioning that any "regular section" only presents approximately 0.3% of the volume of a typical tumor (Simon et al., 2004) and is thus already "biased" significantly.

Several approaches have been used previously to generate results from TMAs with multiple cores of tissue samples that reflect most accurately the expression in the tissue examined (the "case"). Most of these were reported for immunohistochemical studies. In principle, three obvious approaches can be taken: (i) only the areas with a high expression ("hot spots") are selected or (ii) all tissues belonging to the same case are evaluated for their average expression or (iii) a binary scoring system is used where "negative" is defined as "all cores negative" and positive as "at least one core/area is positive" (Kononen et al., 1998; Moch et al., 1999). (i) The first method of choosing "hot spots" for the evaluation was frequently used to assess angiogenesis (CD31) or proliferation markers (Ki-67) (Chiesa-Vottero et al., 2003; Nocito et al., 2001; Rubin et al., 2002), and did show good concordance with the results generated from full tissue sections. (ii) The evaluation of average expression was used in most studies, for example for the evaluation of the Her2/neu (Simon et al., 2001), however the exact evaluation systems used did vary between different studies or were not described in detail and are controversial (Paik, 2003). (iii) The binary system was common for markers like ER and PR expression, similar to the assessment of these for regular sections.

In summary, no standard evaluations system has been established so far for the ISH or IHC. Whether a particular gene is likely to be of significance due to a few "hot spots" of expression or because of broad expression will be different for distinct genes. We did therefore consider several approaches for our TMA studies. During our studies we investigated different genes on sequential

sections of TMA and we will detail the considerations and methods established to evaluate them further below after the "Methods" section.

Methods

1. General protocol for in situ hybridization

Specificity control considerations

The signal-noise ratio of each probe set (antisense and sense) should be established initially on full serial sections of diseased and normal tissue specimen. E.g. to assess expression of a given gene in a breast cancer tissue microarray (TMA) we first use two sets of serial sections from 20 invasive breast cancers and 20 normal breast samples. One set of these samples is hybridized with the antisense probe and one set with the sense probe in one batch experiment. This initial experiment establishes the signal to be expected above a certain background. Also, full sections can later be used as a positive (antisense) and negative (sense) internal control when staining TMAs. With this control included in each experiment, it is reasonable to stain the TMA with an antisense probe only and thus save this precious resource. Typically we will stain two sets of full sections with an sense and antisense probe along with a TMA that is stained with the antisense only.

Generation of probes

A flow chart of the steps is given in in 0: First, a suitable template sequence in the gene of interest is identified. The main criteria for this selection are uniqueness of the selected sequence and appropriate length. The match with other genes or cDNAs can be assessed using e.g the BLAST search engine and the public database of NIH/NCBI (http://www.ncbi.nlm.nih.gov/BLAST) and a <25% match should not cause cross-hybridization. An optimal length of a probe is between 500 and 700 base pairs. Second, to generate the RNA probes by reverse transcription a variety of DNAs can serve as a template. These templates must contain a T7 or Sp6 promoter sequence as a transcription start site. Two approaches can be used to add the T7 or Sp6 promoter sequences (see Fig. 1), (i)

subcloning of a fragment into a bacterial expression vector, which contains the promoters adjacent to the multiple cloning site (e.g. pcDNA3 or others from Invitrogen, Carlsbad, CA) or (ii) PCR with primers containing the T7-sequence in addition to the primers specific to the gene of interest. In general a vector construct is preferable since it will usually render a higher probe yield and is worth the additional effort of subcloning if a large series of experiments is to be performed for that sequence. Once the DNA-template is prepared, Digoxigenin labeled RNA probes are transcribed using the DIG RNA labeling kit (Roche Diagnostics Corp., Indianapolis, IN) according to the manufacturer's protocol.

Preparations and general considerations

The Digoxigenin labeled RNA probes are prepared using the DIG RNA labeling kit (Roche Diagnostics Corp., Indianapolis, IN) according to the manufacturer's protocol. From our experience we recommend a probe generated from 300 to 700 bp of consecutive cDNA sequence which is unique for the mRNA of interest, i.e. does not match any other sequence in a genomic or EST nucleotide-Blast search (see above, NCBI).

Prevention of RNAse contamination is crucial during preparation of the probe and on day 1 of the ISH. The workbench and all bottles, plastic containers, slide holders, beakers and graded cylinders to be used before or on day 1 are cleaned with RNAse AWAY® and left overnight to dry completely. Sterile 15ml or 50ml Sarstedt® or Falcon® tubes should be used for measuring and handling of small solution volumes and the hybridization solution with the RNA probe; they can be considered RNAse free. Most solutions (Table 2) can be prepared in advance. Some must be made on the day of usage and are indicated as such.

Paraffin sections for the ISH should be made by collecting 4 to 5 µm cuts of the paraffin block in a DEPC-H_2O water bath to reduce the risk of RNase exposure due to contaminated water.

Sections are mounted on plus charged slides. In our experience, sections can be stored for weeks to several months in a regular slide container at room temperature without affecting the mRNA quality. Never allow the tissue to dry during the ISH. This is especially important during application of hybridization, antibody and chromate solution.

2. Stepwise day-by-day protocol

(For solutions and materials see Table 2 and 3).

Day 1

Deparaffination and rehydration

Place slides in the slide holder, incubate for 1 hour at 65°C then deparaffinize with two Xylene incubations, 5 min each. Wash twice with Xylene (5 min) and rehydrate with Ethanol 100% twice, 5 min each, followed by DEPC-H_2O once, 5min.

Protein digestion

The slides are transferred to PBS at 37°C for 5 min and tissue proteins are then digested by incubation with 10 µg/ml Proteinase K/PBS at 37°C for 10 min. Enzyme is washed out by retransferring the slides to PBS 37°C for 5 min.

HCl treatment and Acetylation

Slides are washed in DEPC-H_2O and SSC 2x/DEPC-H_2O for 5 min each then transferred to 0.2M HCl for 15 min. *(During these 15 min the Digoxigenin labeled probes are collected from the storage at -80°C and placed on ice for slow thawing, then the 0.25% acetic anhydride solution is mixed. The incubation chambers are also prepared by placing filter paper on the bottom, moistening it with DEPC-H2O and covering it with Parafilm®.)* After the 15 min HCl incubation, slides are

washed in 0.1M TEA-HCl pH 8.0 for 5 min, then acetylated with the 0.25% acetic anhydride / 0,1M TEA-HCl pH 8,0 solution for 15 min.

Preparation of the probe/hybridization solution

The Digoxigenin probes are mixed with the hybridization solution while the slides are acetylated. We recommend calculating 400 µl of hybridization solution per TMA slide. After estimating and aliquotting the necessary total volume of solution for each probe into an RNase free tube, the Digoxigenin probe is added to a final concentration of 1,500 ng probe per 1 ml solution and the mixed solution is stored on ice. *(Example for a series of 5 TMA slides and 4 regular sections for external controls. Antisense: Mix 3,000 µl of hybridization solution (5x400+4x250) and 4,500 ng of antisense-RNA-probe. Sense: Mix 1,000 µl of hybridization solution (4x250) and 1,500 ng of sense-RNA-probe.)*

Hybridization

After acetylation, slides are placed in SSC 2X. After 5 min the slides are then processed individually as following: The back of each slide and the front areas without tissues are carefully dried with a Kimwipe® and the slide is placed horizontally on the Parafilm® in the hybridization chamber. Enough probe/hybridization solution mix to cover all tissue-areas is now carefully applied to the slide with a pipette (usually 300-400 µl for a TMA). The next slide is now treated until all slides are covered with solution. The chamber is closed, sealed with Parafilm® and incubated O/N at 42°C. *(Note: Some ISH-protocol recommend calculating the "optimal temperature" and/or testing different temperatures. When using our relatively large probes, we observed a good sensitivity and specificity at 42°C and did not have to optimize the temperature)*

Day 2

Post-hybridization wash

The hybridization solution is washed off each slide with 1 to 2ml of SSC 2X using a plastic Pasteur pipette and the slide is then placed in a slide holder in SSC 2X. Once all slides are collected, they are first transferred to fresh SSC 2X for 5 min (room temperature) then washed twice in SSC 2X at 52°C for 10 min each.

RNA digestion

It is crucial to remove unbound probe to reduce unspecific background: Slides are placed in STE buffer at 37°C for 5 min then treated with a solution of 2 mg RNase A in 200 ml STE buffer at 37°C for 10 min. Washout is performed by STE buffer at 37°C for 5 min.

Refixation

Slides are incubated with SSC 2x/Formamide at 42°C for 10 min for fixation and crosslinking and then washed with SSC 1X and SSC 0.5X at 42°C for 5 min each.

Blocking and antibody incubation

After washing the slides for 1 min in buffer #1 (room temperature), blocking is performed with 2% horse serum/buffer #1 for 30 min. *(Note: Bovine serum does contain alkaline phosphatase and should not be used.)* Slides are individually taken out of the blocking solution and the back and tissue free areas of the front are dried with Kimwipes®. The slide is placed in the hybridization chamber and 500 to 700 µl of antibody solution are pipetted onto the tissue area with a plastic Pasteur pipette. The chamber is sealed, once all slides are processed, and placed horizontally in a refrigerator at 4°C for O/N incubation.

Day 3

Preparation of the staining solution

The staining solution is prepared by dissolving NBP/BCIP tablets in pure H_2O (one tablet/10 ml H_2O). 1.5 to 2ml solution will be used per slide. The solution is light sensitive, therefore the tube is wrapped in aluminum foil and stored in a dark drawer. Dissolving is slow and can be enhanced by occasional vortexing.

Antibody washout

The antibody solution is flushed off each slide with 1 to 2ml of buffer #1 using a plastic Pasteur pipette. The slide is then placed in a slide holder in buffer #1. Once all slides are collected they are first transferred twice to fresh buffer #1 for 5 min each then twice to buffer #2 for 5 min each.

Staining

One slide at a time is taken out of the buffer #2 and the back and tissue free front areas are dried as detailed before. The slide is placed in the chamber and 1 to 2ml of staining solution are applied with a plastic Pasteur pipette. Once all slides are processed, the chamber is placed in a dark drawer until sufficient staining is observed. The time required for developing can range from 30 minutes to several hours. Checking the staining progress of individual slides under a microscope (after which fresh staining solution should be reapplied) can be used to monitor the process.

Terminating the staining and washing

When sufficient staining is observed, slides are washed twice in buffer #2 to remove the NBT/BCIP solution and then transferred to buffer #3 for 10 min. Slides are then washed in 0.5% Tween 20 on a rocker or vertical shaker for 5 min after which they are transferred to a H_2O bucket. The H_2O bucket is carefully flushed under a ddH_2O tub until the Tween is washed out.

Mounting

The slides are placed with their backs down on a filter paper for at least 1 hr to allow complete drying. The dried slides can then be mounted with Cytoseal™ and cover slips.

Results and Discussion

Appearance of ISH staining

Positive staining results will appear as a violet to brown cytoplasmic staining and nuclei appear light or unstained (0a). A granulated staining pattern can often be observed in the cytoplasm when the probe targets mitochondrial RNA rather than a cytosolic mRNA. Unspecific staining in collagen structures does sometimes occur but can easily be distinguished due to its blue rather than violet color and the lack of cells and unstained nuclei. Counterstaining with e.g. Hematoxylin is not necessary and is not recommended since the comparable low staining intensity of the ISH may otherwise be difficult to evaluate. In our experience, tissue structures can still be determined, even in completely negative or sense control tissues (0b). Although a normal light microscope is used to evaluate the ISH, light intensities should be set lower, compared to the microscopy of standard stainings like H&E or immunohistochemistry.

Evaluation of staining for TMA cores or tissues

Staining for each TMA core or full tissue section should always be evaluated in conjunction with the sense controls of that batch. We use a semi quantitative four-tier system: The highest staining intensities and frequencies are defined as "+++", low but certain staining "+" and intensities, which are average to high "++". The details of the criteria to evaluate staining results of the individual cores are shown in Table 4. An example of stainigs of a TMA is shown in 0 (ISH for PTN).

Although we have found the ISH to be a reliable and stable method, a variation of staining intensities may often be observed inter- and intra-experimentally due to variations in the experimental conditions or different mRNA preservation. Therefore, when using TMAs, each of

them is evaluated individually as follows: First the highest, lowest and average staining intensities are identified individually for each slide and then used to determine the criteria for "negative", "low", "medium" and "high positive" for that particular TMA. This is necessary, since we sometimes observe that results did vary in the maximum and average staining between different TMAs, in spite of having processed them with the same probe and in the same batch. However, when adjusting the criteria as detailed, we found that the distributions of the staining results from the individual arrays were identical (all $p>0.8$ by Mann-Whitney-test). Also, the correlations with biological parameters (including metastasis-status; estrogen-, progesterone-, HER2-receptor-status; stage; grade and survival time) were similar in the statistical analyses, suggesting that the individual calibration of the evaluation criteria for each array is feasible and useful.

Depending on the hypothesis to be tested, further evaluation of the TMA ISH data may vary. In principle, since usually several individual cores will present a given donor tissue, we have to distinguish between results from individual "cores", where each core contributes a data point and "case" results, where each donor tissue represents one data point derived from all of the cores representing that case.

Obtaining case results from TMA cores

We tested different approaches to generate representative case results, since multiple cores typically represent each case. Initially we classified results according to the percentage of stained cancer cells over all cores of each case as: negative (<5%), focal (5-25%), positive (25-75%), high positive (75%-90%) and very high positive (>90%). Although similar systems have been reported in the literature and these case values did show results similar in most statistical analyses to those for the cores and are thus valid in general, we refined this approach due to the following reasons:

(i) the information of the staining intensity, which is included in the core results, was lost;

(ii) the comparison with the results of full sections and reference tissues on the TMA, which are often only present as one core, was difficult due to the discrepancy in the evaluation criteria and

(iii) if a TMA is evaluated in a double blind (i.e. it is not known which cores do belong to the same case) it is impossible to grade the case with the percentage system.

Therefore we did test a different approach, taking advantage of one of the main features of TMAs as pointed out previously (Bubendorf et al., 2001), namely the fact that the large number of observations generated by a TMA can be assumed to correct an error in the assessment of an individual case. We established an unbiased mathematical approach to generate case values from the cores:

First, we assigned each core a discrete value for the staining results ranging from 1 (for "-") to 4 (for "+++). Consecutively we assumed that (i) the average of all cores values of a case would be representative for the mRNA expression of that case, (ii) that reassigning the averages of all cases to the same four categories (1-4) could be performed with an appropriate cut-off point and (iii) that the re-categorized values would have to show similar results in the subsequent analysis as the percentage values and the core values. The appropriate cut-off points were chosen for each study by comparing the initial distribution of the core values and the resulting case value distribution aiming for the closest possible match. An example is given in 0.

To validate our above stated hypotheses we compared the correlation of *a)* the core values, *b)* the case values from the percentage classification system and *c)* the average system with other parameters. In summary, the average system showed similar significant correlations or trends with biological parameters as the core values did (data not shown). The percentage classification results were mostly comparable but often showed less and sometimes no significance. In one of our studies we investigated breast cancer progression tissue microarrays from the CBCTR (NCI/NIH, Bethesda,

MD) for the expression of several genes and the protein expression of one of these. We found, after double blind evaluation, that the core values and the average system values could both distinguish between the three groups of breast cancers, present on the TMAs (node negative tumors, node positive tumors and tumors with distant metastases).

We believe that the arithmetic mean with consecutive reclassification is a valid approach to address the relevance of most mRNAs by ISH on TMAs with multiple cores per case. However, there is a caveat: For those genes where "hot spots" are assumed to be crucial, it might be feasible to either focus on areas with maximum expression or to choose the core with maximum staining to represent the case.

Correlation of Protein and mRNA

In one of our studies we detected AIB1 mRNA in 45% of low-grade pancreatic neoplasia in situ (PanIN) by in situ hybridization but the protein was only detectable in 27% (Immunohistochemistry) (Henke et al., submitted). With progression to pancreatic adenocarcinoma, the mRNA expression (74% positive) and protein expression (64% positive) did converge and the comparison of the mRNA and protein expression levels did not show any significant differences. This co-analysis off mRNA and protein expression did thus provide additional information consistent with the idea of AIB1 being increasingly expressed throughout advancing stages of pancreatic malignant progression. It also did show that the mRNA in situ hybridization, detailed in this review, has a high sensitivity. It detected the mRNA of AIB1 in >99.9% of the samples which had a protein staining of "++" or "+++" and still in >85% of those being low positive ("+") for the protein. In addition, we detected mRNA in a few protein-negative cases especially in the low grade PanIN.

Further analysis of ISH results in relation to other data (biological, medical history, protein or other mRNAs)

When addressing the question of correlation of expression between an mRNA and its protein product or between different mRNAs, it is appropriate to use the results from cores for each mRNA and protein if they originate from parallel sections of the same TMA block(s). This is feasible since this type of correlation is truly a function of staining of tissue and cells and each core can therefore be seen as a distinct observation, similar to choosing different areas on a regular tissue section. To illustrate this, we stained six sets of parallel TMA sections, each set for a different mRNA. All genes were hypothesized to have a positive correlation with tumor metastasis. Using the "core results" of all six TMAs for a Spearman non-parametric correlation, we observed that expression of all genes did correlate well (all $r>0.5$, all $p<0.001$, all $n>80$) suggeting that all share a correlation with tumor metastasis.

For other evaluations, especially correlation with biological or clinical parameters, it usually is not appropriate to only use the results from cores alone and thus results for cases should be generated. Nevertheless, we recommend evaluating the "core results" in parallel or prior to the "case results" for several reasons:

1. The analysis of the core results with regard to correlation with other parameters is easy as no additional steps to generate case results have to be performed and can often be used to determine if a further analysis is justified.
2. Correlations with biological parameters might be statistically significant for the core results but may sometimes only show a borderline trend for the case data. This either indicates that no true correlation exists or, as we have frequently observed, that the case number was too low to achieve significance. A new study with more cases is then needed.

3. A discrepancy between core and case findings could also indicate that the method to define the case results might not be feasible to address the question. Choosing a "hot spot" or a binary classification system rather then the "standard" average based system might clarify this. Therefore, all these systems should be consecutively evaluated in parallel to the core analysis.

Conclusions

The ISH with Digoxigenin labeled RNA probes is an established method to address the expression of a target RNA in tissues. We have optimized this method for the use with formalin-fixed paraffin-embedded archival tissues using regular sections and TMAs with single or multiple sample cores per case. When using large probes at a high concentration, we achieved good sensitivity while retaining excellent specificity for all of the RNAs investigated and found the ISH to be a very reliable and stable method, inspite of a frequently voiced concerns regarding variable mRNA quality in paraffin sections. We did establish an unbiased evaluation system for TMAs for genes for which a broad expression is assumed to be relevant. This system was valid in our analyses and can be easily adapted to evaluate genes for which "hot spot" expression rather than average expression is critical. Our findings justify the use of the ISH to study the expression of genes in large sample numbers either alone or in conjunction with protein analyses. Our reported method enables investigators to screen such large populations with low experimental effort when using TMAs. We propose that ISH can be used to assess the significance of a gene of interest, before the respective protein can be investigated. The ISH can thus maximize the use of available time and resources by distinguishing early between more and less relevant candidate genes while already providing crucial data to correlate gene expression with biological parameters.

Acknowledgements

We wish to thank Dr. Sheila Taube, National Cancer Institute, Bethesda and Dr. Baljit Singh, Georgetown University for their input and valuable discussions as well as Dr. Kenneth Pienta, U. Michigan, Ann Arbor for kindly providing tissue microarrays containing prostate tissue samples. Other tissue microarrays were obtained from the CBCTR of the National Cancer Institute / NIH, from Georgetown University and from Johns Hopkins University. This work was supported in parts by a grant of the Department of Education and Research of the Federal Government of Germany (to R.T.H.), by a Johns Hopkins Clinical Scientist Award (to A.M.) and by grants of the NIH as well as the DOD Breast Cancer Program (to A.W.). The Johns Hopkins tissue microarray facility is supported by a generous grant from the family of Margaret Lee.

References

Bismar, T. A., Bianco, F. J., Zhang, H., Li, X., Sarkar, F. H., Sakr, W. A., Grignon, D. J., and Che, M. (2003). Quantification of G250 mRNA expression in renal epithelial neoplasms by real-time reverse transcription-PCR of dissected tissue from paraffin sections. Pathology *35*, 513-517.

Bubendorf, L., Nocito, A., Moch, H., and Sauter, G. (2001). Tissue microarray (TMA) technology: miniaturized pathology archives for high-throughput in situ studies. J Pathol *195*, 72-79.

Camp, R. L., Charette, L. A., and Rimm, D. L. (2000). Validation of tissue microarray technology in breast carcinoma. Lab Invest *80*, 1943-1949.

Chiesa-Vottero, A. G., Rybicki, L. A., and Prayson, R. A. (2003). Comparison of proliferation indices in glioblastoma multiforme by whole tissue section vs tissue microarray. Am J Clin Pathol *120*, 902-908.

Duyster, J., Bai, R. Y., and Morris, S. W. (2001). Translocations involving anaplastic lymphoma kinase (ALK). Oncogene *20*, 5623-5637.

Fend, F., and Raffeld, M. (2000). Laser capture microdissection in pathology. J Clin Pathol *53*, 666-672.

Fink, L., Seeger, W., Ermert, L., Hanze, J., Stahl, U., Grimminger, F., Kummer, W., and Bohle, R. M. (1998). Real-time quantitative RT-PCR after laser-assisted cell picking. Nat Med *4*, 1329-1333.

Jin, L., Thompson, C. A., Qian, X., Kuecker, S. J., Kulig, E., and Lloyd, R. V. (1999). Analysis of anterior pituitary hormone mRNA expression in immunophenotypically characterized single cells after laser capture microdissection. Lab Invest *79*, 511-512.

Kabbarah, O., Pinto, K., Mutch, D. G., and Goodfellow, P. J. (2003). Expression profiling of mouse endometrial cancers microdissected from ethanol-fixed, paraffin-embedded tissues. Am J Pathol *162*, 755-762.

Kagan, B. L., Henke, R. T., Cabal-Manzano, R., Stoica, G. E., Nguyen, Q., Wellstein, A., and Riegel, A. T. (2003). Complex regulation of the fibroblast growth factor-binding protein in MDA-MB-468 breast cancer cells by CCAAT/enhancer-binding protein beta. Cancer Res *63*, 1696-1705.

Klomp, H. J., Zernial, O., Flachmann, S., Wellstein, A., and Juhl, H. (2002). Significance of the expression of the growth factor pleiotrophin in pancreatic cancer patients. Clin Cancer Res *8*, 823-827.

Kononen, J., Bubendorf, L., Kallioniemi, A., Barlund, M., Schraml, P., Leighton, S., Torhorst, J., Mihatsch, M. J., Sauter, G., and Kallioniemi, O. P. (1998). Tissue microarrays for high-throughput molecular profiling of tumor specimens. Nat Med *4*, 844-847.

Kurtz, A., Wang, H. L., Darwiche, N., Harris, V., and Wellstein, A. (1997). Expression of a binding protein for FGF is associated with epithelial development and skin carcinogenesis. Oncogene *14*, 2671-2681.

List, H. J., Reiter, R., Singh, B., Wellstein, A., and Riegel, A. T. (2001). Expression of the nuclear coactivator AIB1 in normal and malignant breast tissue. Breast Cancer Res Treat *68*, 21-28.

Liu, X. H., Aigner, A., Wellstein, A., and Ray, P. E. (2001). Up-regulation of a fibroblast growth factor binding protein in children with renal diseases. Kidney Int *59*, 1717-1728.

Mashour, G. A., Wang, H. L., Cabal-Manzano, R., Wellstein, A., Martuza, R. L., and Kurtz, A. (1999). Aberrant cutaneous expression of the angiogenic factor midkine is associated with neurofibromatosis type-1. J Invest Dermatol *113*, 398-402.

Moch, H., Schraml, P., Bubendorf, L., Mirlacher, M., Kononen, J., Gasser, T., Mihatsch, M. J., Kallioniemi, O. P., and Sauter, G. (1999). High-throughput tissue microarray analysis to evaluate genes uncovered by cDNA microarray screening in renal cell carcinoma. Am J Pathol *154*, 981-986.

NCBI. BLAST. http://www.ncbi.nlm.nih.gov/BLAST/

Nocito, A., Bubendorf, L., Maria Tinner, E., Suess, K., Wagner, U., Forster, T., Kononen, J., Fijan, A., Bruderer, J., Schmid, U., *et al.* (2001). Microarrays of bladder cancer tissue are highly representative of proliferation index and histological grade. J Pathol *194*, 349-357.

Paik, S. (2003). Clinical trial methods to discover and validate predictive markers for treatment response in cancer. Biotechnol Annu Rev *9*, 259-267.

Panoskaltsis-Mortari, A., and Bucy, R. P. (1995). In situ hybridization with digoxigenin-labeled RNA probes: facts and artifacts. Biotechniques *18*, 300-307.

Powers, C., Aigner, A., Stoica, G. E., McDonnell, K., and Wellstein, A. (2002). Pleiotrophin signaling through anaplastic lymphoma kinase is rate-limiting for glioblastoma growth. J Biol Chem *277*, 14153-14158.

Ray, R., Cabal-Manzano, R., Moser, A. R., Waldman, T., Zipper, L. M., Aigner, A., Byers, S. W., Riegel, A. T., and Wellstein, A. (2003). Up-regulation of fibroblast growth factor-binding protein, by beta-catenin during colon carcinogenesis. Cancer Res *63*, 8085-8089.

Rubin, M. A., Dunn, R., Strawderman, M., and Pienta, K. J. (2002). Tissue microarray sampling strategy for prostate cancer biomarker analysis. Am J Surg Pathol *26*, 312-319.

Schulte, A. M., Malerczyk, C., Cabal-Manzano, R., Gajarsa, J. J., List, H. J., Riegel, A. T., and Wellstein, A. (2000). Influence of the human endogenous retrovirus-like element HERV-E.PTN on the expression of growth factor pleiotrophin: a critical role of a retroviral Sp1-binding site. Oncogene *19*, 3988-3998.

Simon, R., Mirlacher, M., and Sauter, G. (2004). Tissue microarrays. Biotechniques *36*, 98-105.

Simon, R., Nocito, A., Hubscher, T., Bucher, C., Torhorst, J., Schraml, P., Bubendorf, L., Mihatsch, M. M., Moch, H., Wilber, K., *et al.* (2001). Patterns of her-2/neu amplification and overexpression in primary and metastatic breast cancer. J Natl Cancer Inst *93*, 1141-1146.

Steel, J. H., Hamid, Q., Van Noorden, S., Jones, P., Denny, P., Burrin, J., Legon, S., Bloom, S. R., and Polak, J. M. (1988). Combined use of in situ hybridisation and immunocytochemistry for the investigation of prolactin gene expression in immature, pubertal, pregnant, lactating and ovariectomised rats. Histochemistry *89*, 75-80.

Steel, J. H., Jeffery, R. E., Longcroft, J. M., Rogers, L. A., and Poulsom, R. (1998). Comparison of isotopic and non-isotopic labelling for in situ hybridisation of various mRNA targets with cRNA probes. Eur J Histochem *42*, 143-150.

Steel, J. H., Morgan, D. E., and Poulsom, R. (2001). Advantages of in situ hybridisation over direct or indirect in situ reverse transcriptase-polymerase chain reaction for localisation of galanin mRNA expression in rat small intestine and pituitary. Histochem J *33*, 201-211.

Van Deerlin, V. M., Gill, L. H., and Nelson, P. T. (2002). Optimizing gene expression analysis in archival brain tissue. Neurochem Res *27*, 993-1003.

Table 1. Overview of probes and tissues used with the ISH protocol described.

Gene/cDNA		probe length	Normal and neoplastic tissues analyzed	protein analyzed	Reference
h Pleiotrophin		551	BC(T), CC(T), PC(T), GBM, BileC(T), AC (T)	Yes	1
Anaplastic lymphoma kinase		648 (*)	BC(T), CC(T), PC(T), GBM, BileC(T), AC (T)	No	2
Midkine		>500	Neurofibroma, Skin	Yes	3
m Pleiotrophin		547	mouse mammary glands	No	Unpublished
h fibroblast growth factor binding protein 1		>500	BC(T), CC(T), PC(T), AC (T)	Yes	4
m fibroblast growth factor binding protein 1		>500	mouse colon, colon adenoma and skin	Yes	5
amplified in breast cancer 1 (AIB1)		679	BC(T), PC(T), PanIN(T), ProstateC	Yes	6, 7
transgene mRNA: Δ3-AIB1/pcDNA3		300 (**)	mouse mammary glands (normal mice, and transgenes for Δ3-AIB1/pcDNA3)	No	8
LS42	metastasis homing genes	656	BC(T), CC (T), Insulinoma(T), ProstateC(T)	No	Unpublished
JN40		874	"	No	Unpublished
SW48		429	"	No	Unpublished
SW50		750	"	No	Unpublished
LS45		715	"	No	Unpublished

h = human; m = mouse; (T); BC: normal breast and breast cancer; CC: normal colon and colon cancer (primary and metastases); PC: normal pancreas and pancreatic adenocarcinoma; PanIN: normal pancreas and pancreatic intraepithelial neoplasia; GBM: Glioblastoma multiforme; BileC: bile, liver and bile duct Cancer; AC: intestine, pancreas and ampullary carcinoma; ProstateC: normal prostate and prostate cancer (primary and metastases); (*): The probe for detection of anaplastic lymphoma kinase was directed against the coding region of the extracellular ligand-binding domain and does not hybridize with translocation products like NPM-ALK (Duyster et al., 2001). (**): To investigate expression of a human Δ3-AIB1 transgene, introduced into nude mice, we designed a probe for the human-specific 3'UTR (140 bp) and the consecutive 160 bp from the expression vector used to introduce the gene, resulting in a 300 bp probe. References: 1, Klomp et al., 2002; 2, Powers et al., 2002; 3, Mashour et al., 1999; 4, Kagan et al., 2003; 5, Ray et al., 2003; 6, List et al., 2001; 7, Henke et al., submitted; 8, Tilli et al., under revision

Table 2. Solutions required for the ISH

Day 1 (RNAse free solutions)	**1)**	Digoxigenin labeled RNA-Probe / approx. 600 ng per slide (microarray)
	2)	Xylene (RNA-grade)
	3)	Ethanol 100% (RNA-grade)
	4)	DEPC-H_2O [add 1ml Diethyl pyrocarbonate per 1l ddH_2O, stir overnight and autoclave]
	5)	PBS 10x [80g NaCl, 2g KCl, 11.5g Na_2HPO_4*7H_2O, 2g KH_2PO_4 ad 1l DEPC-H_2O]
	6)	PBS 1x [100ml PBS 10x, 900ml DEPC-H_2O]
	7)	Proteinase K / PBS 1x [prepare on Day 1: 2mg Proteinase K (Roche Diagnostics GmbH, Germany) in 200ml PBS 1x (final conc. 10 µg/ml)]
	8)	SSC 20x [3M NaCl, 0,3M Na-citrate in DEPC-H_2O; alternatively: commercial stock solution]
	9)	SSC 2x [100ml SSC 20x, 900ml DEPC-H_2O]
	10)	0,2M HCl [1 Part 6N HCl, 29 Parts DEPC-H_2O]
	11)	TEA-HCl [dissolve 18,6g Triethanolamine-HCl in 900ml DEPC-H_2O, titrate to pH 8,0 with 10N NaOH, ad 1l DEPC-H_2O]
	12)	0,25% Acetic anhydride/TEA-HCl [prepare on Day 1: 0.5ml Acetic anhydride in 200ml TEA-HCl]
	13)	Hybridization solution [Sigma-Aldrich, St. Louis, MO, USA, Cat-Nr.: H7782; 400 µl/TMA]
Day 2 and 3 (normal solutions)	**14)**	dd-H_2O
	15)	STE buffer [500mM NaCl, 20mM Tris-HCl pH 7.5, 1mM EDTA in dd-H_2O]
	16)	RNase A / STE buffer [prepare on Day 2: 2g of RNase A in 200ml STE Buffer]
	17)	SSC 2x [200ml SSC 20x, ad 2 liters dd-H_2O]
	18)	SSC 2x / Formamide [prepare on Day 2: 100ml Formamide, 100ml SSC 2x]
	19)	SSC 1x [10ml SSC 20x, 190ml dd-H_2O]
	20)	SSC 0,5x [5ml SSC 20x, 195ml dd-H_2O]
	21)	Buffer #1 [100mM Tris-HCl pH 7.5, 150mM NaCl in dd-H_2O, FILTER]
	22)	2% Horse Serum / buffer #1 [prepare on Day 2: 4ml of Horse Serum in 196ml buffer #1]
	23)	antibody solution (calculate 500-700 µl/slide) [prepare on Day 2: dilute in buffer #1: Anti-Digoxigenin-AP Fab fragments (Roche Diagnostics GmbH, Cat-Nr: 1 093 274) 1:250, Horse Serum 1:100]
	24)	Buffer #2 [100mM Tris-HCl pH 9.5, 100mM NaCl, 50mM $MgCl_2$ in dd-H_2O, FILTER]
	25)	NBT/BCIP Solution (calculate 500-1000µl / Tissue; **light sensitive,** wrap in aluminum-foil) [prepare on Day 2: One FAST BCIP/NBT Tablet (Sigma-Aldrich) per 10ml distilled H_2O] [alternatively: dilute in buffer #2: 33,75µl / 10ml (3,75mg / 10ml) of NBT (nitroblue tetrazolium) and 35µl / 10ml (1,75mg / 10ml) of BCIP (toluidinum salt)]
	26)	Buffer #3 [10mM Tris-HCl pH 8,0, 1mM EDTA in dd-H_2O; FILTER]
	27)	0,5% Tween 20 [1ml Tween 20 in 200ml dd-H_2O]

Table 3. Materials and equipment required for the ISH

1)	RNAse AWAY® (Molecular BioProducts Inc., San Diego, CA, USA)
2)	200ml plastic containers for the ISH solutions
3)	Plastic slide holders for 24 slides
4)	Plastic trays to hold slides vertically during the o/n hybridization, the o/n antibody incubation and the NBT/BCIP staining
5)	Kimwipes® (Kimberly-Clark Inc., Mississauga, Ontario, CA)
6)	Parafilm® (Pechiney Plastic Packaging, Menasha, WI, USA)
7)	Plastic – Pasteur pipettes
8)	Vacuum system for at least 500 – 1000ml of volume
9)	Empty drawer to develop during NBT/BCIP staining
10)	Workbench treated with RNAse AWAY®
11)	Vacuum filter system for buffer #1, #2 and #3 (Corning Inc., Corning, NY, USA)
12)	Ice
13)	Water bath at 37°C and 42°C
14)	Oven with adjustable temperatures (65°C and 52°C)
15)	Incubator at 42°C for o/n hybridization
16)	Mounting medium: Cytoseal 60, low viscosity (Stephens Scientific, Kalamazoo, MI, USA)
17)	Cover slips for microscopical slides

Table 4. Criteria to evaluate the ISH staining of a TMA core or full tissue section.

Negative staining	-	no staining observed in this core/section that exceeds the sense control signals
Positive staining	+	positive staining exceeding sense control signals in at least 20% of the cells
	++	above average staining intensity in all cells OR high staining in at least 50% of the cells and "+" intensity in the remainder
	+++	high staining intensity in all cells OR very high staining in at least 50% of the cells and "++" intensity in the remainder
N/A	X	core missing or the tissue of this core is necrotic or autolysed and the RNA is degraded
	-NT	the appropriate cells (for example cancer cells in a carcinoma core) are not present and no staining is observed in other cells.
	+NT	the appropriate cells (for example cancer cells in a carcinoma core) are not present but staining is observed in other cells. *Note: if the positive celltype can be identified it is feasible to change the tissuetype associated with that core and to evaluate it as "+, ++ or +++" for that appropriate tissuetype instead of "+NT".*

The system is optimized to assess a gene for which the average expression pattern is hypothesized to be relevant. If "hot spots" are assumed to be critical it should be adapted as needed.

Legends to the Figures

Figure 1. Flowchart: Generating a Digoxigenin-labeled-RNA probe from a DNA template.

Transcription-templates contain a T7 or Sp6 promoter sequence to be transcribed (bottom box). This can be generated in two distinct ways outlined on the left and right respectively.

Left: The desired sequence (5'-gene-3') is subcloned into a suitable bacterial expression vector with consecutive linearization before OR after insertion. Each linearized plasmid then serves as the template for the antisense OR the sense probe.

Right: A two-step PCR is performed from any suitable template containing the desired sequence. That sequence is first amplified with sequence-specific primers. The product is then gel-purified away from primers and used as a template for the second step. During the second PCR, either the forward OR the reverse primers are substituted. The replacement primer contains a short 5' non-specific sequence followed by the T7-sequence (bold-italic) and the primer-sequence that was used in the first PCR, at the 3' terminus.

Figure 2. Representative ISH staining results for PTN-antisense and -sense probe staining.

Two adjacent parallel 4 μm sections of a paraffin embedded pancreatic adenocarcinoma were hybridized with the antisense (A) or sense (B) Digoxigenin labeled probe for Pleiotrophin (40X image is shown)

A) The antisense probe shows a strong cytoplasmic staining in cancer cells (arrows). Nuclei appear unstained ("empty" *) or significantly less stained (#) if they are cut tangentially and not through their center. The color may vary from a pink-violet (low intensity) over blue-violet (as shown) to an intense, indigo or brown color, depending on the intensity of the staining and the exact ratio of NBT to BCIP applied to develop the ISH.

B) The sense RNA probe (negative control) shows no signal. Note that tissue structures can still be identified in the sense control, although the slides were not counterstained. In particular, the cancer cell clusters can be easily distinguished from the tumor stroma (X).

Figure 3. Example of a TMA stained for Pleiotrophin mRNA by ISH.

A) Cores from two different primary colon carcinoma (left), a normal colon (top right) and a prostate hyperplasia (bottom right) are shown (4X). Cytoplasmic ISH staining was scored in each core for carcinoma cells (left cores), or epithelial cells, respectively. Scores (- to +++) are shown adjacent to the core. **B)** Colon carcinoma. 40x magnification of the area marked in panel A. Strong cytoplasmic staining with "empty nuclei" is found in cancer cells whereas no signals are observed in the tumor stroma.

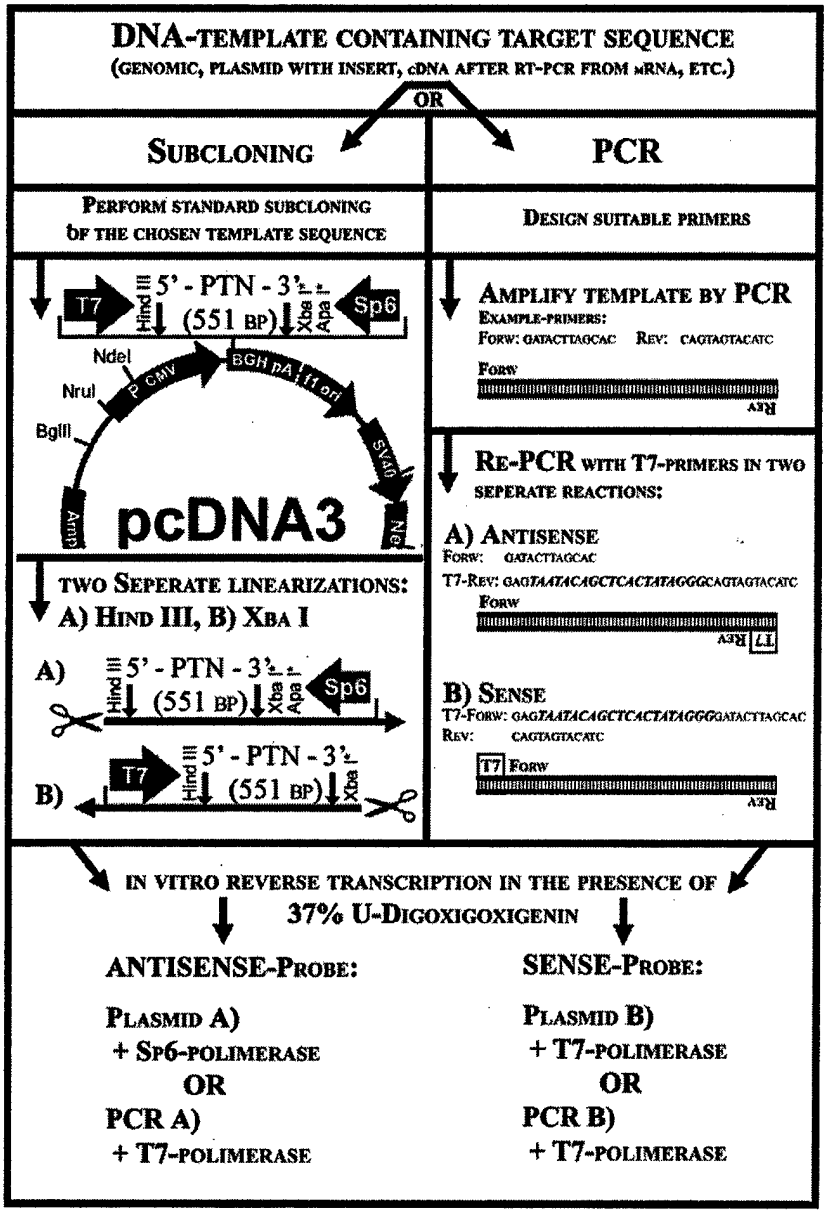
Figure 4. Pleiotrophin expression in ampullary carcinoma. ISH evaluation of replicate TMA samples from different cases.

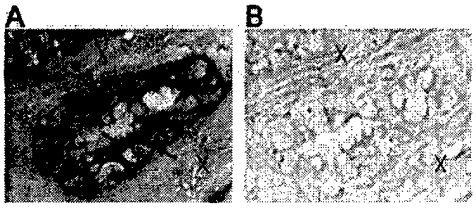
Ampullary carcinoma TMAs 121, 125 and 126 were stained for pleiotrophin by ISH and each core was evaluated. Each TMA contained samples of 18 different tumors represented by four cores.

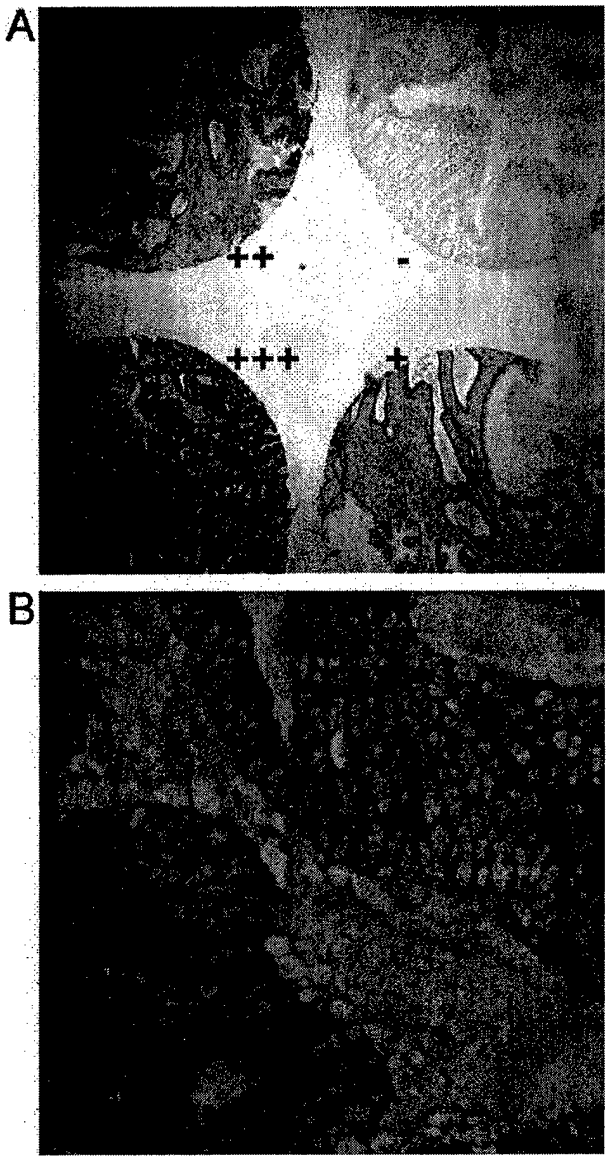
A) Examples for staining evaluation of cores (A-D) arithmetic mean (mean) and assigned case values are shown in detail for 18 of the 54 cases. Core results ranged from 1: "-"; 2: "+"; 3: "++" to 4: "+++". Missing values: "N/A". Each line represents one original tumor (TMA-case). The arithmetic mean of the core values is calculated for each and classified as a distinct value (assigned value) of 1, 2, 3 or 4 using cut-off points.

B) Summary of the core and case values. The cut-off points shown resulted in an optimal match between the distribution of the core and case values. The ratio of all cores to all cases was set to 1 and used to normalize the ratios for each category (relative ratio). A relative ratio of 1.00 would thus reflect a perfect match. Optimal cut-off points are defined as those, where the resulting relative ratios are closest to 1.00, special emphasis is given to the ratios representing the most data-points (here categories 2 and 3).

C) Bar graph of the percentage of cores and cases at different expression levels. Each set of columns represents the percent of values within the same category (1(-) to 4(+++)). Note that both groups (cores and cases) show an almost identical distribution.


DNA-TEMPLATE CONTAINING TARGET SEQUENCE
(GENOMIC, PLASMID WITH INSERT, cDNA AFTER RT-PCR FROM mRNA, ETC.)
OR
SUBCLONING
PCR
PERFORM STANDARD SUBCLONING OF THE CHOSEN TEMPLATE SEQUENCE
DESIGN SUITABLE PRIMERS
T7
5' - PTN - 3'
Hind III
(551 BP)
Xba I
Apa I
Sp6
AMPLIFY TEMPLATE BY PCR
EXAMPLE-PRIMERS:
FORW: GATACTTAGCAC REV: CAGTAGTACATC
FORW
REV
NdeI
NruI
BglII
P CMV
BGH pA
f1 ori
SV40
Amp
pcDNA3
Neo
RE-PCR WITH T7-PRIMERS IN TWO SEPERATE REACTIONS:
A) ANTISENSE
FORW: GATACTTAGCAC
T7-REV: GAGTAATACAGCTCACTATAGGGCAGTAGTACATC
FORW
REV
T7
B) SENSE
T7-FORW: GAGTAATACAGCTCACTATAGGGGATACTTAGCAC
REV: CAGTAGTACATC
T7 FORW
REV
TWO SEPERATE LINEARIZATIONS:
A) HIND III, B) XBA I
A)
B)
IN VITRO REVERSE TRANSCRIPTION IN THE PRESENCE OF
37% U-DIGOXIGOXIGENIN
ANTISENSE-PROBE:
PLASMID A)
+ SP6-POLIMERASE
OR
PCR A)
+ T7-POLIMERASE
SENSE-PROBE:
PLASMID B)
+ T7-POLIMERASE
OR
PCR B)
+ T7-POLIMERASE


A
B


A
++
-
+++
+
B

A

TMA-case	Core A	B	C	D	Mean	assign. Value
121001	3	2	2	2	2,25	2
121002	2	3	2	2	2,25	2
121003	3	2	3	2	2,50	3
004 - 015	⋮	⋮	⋮	⋮	⋮	⋮
121016	3	2	2	2	2,25	2
121017	3	2	3	3	2,75	3
121018	3		4	3	3,33	3
125001	4	3	4	4	3,75	4
125002	2		3	2	2,33	2
125003	2	1		1	1,33	1
004 - 015	⋮	⋮	⋮	⋮	⋮	⋮
125016	1		3	3	2,33	2
125017	1	1	2	3	1,75	2
125018	3	2	3	3	2,75	3
126001	3	4	3	4	3,50	4
126002	3	3	3	2	2,75	3
126003	3	3	2	3	2,75	3
004 - 015	⋮	⋮	⋮	⋮	⋮	⋮
126016	3	3	3	3	3,00	3
126017	3		2	2	2,33	2
126018	3	3	4	3	3,25	3

B

Value	Cores	Cases	cut-off	rel. ratio
1 (neg)	16	5	< 1.75	0.86
2 (+)	73	18	1.75 - 2.5	1.10
3 (++)	92	25	2.5 - 3.4	0.99
4 (+++)	19	6	> 3.4	0.86
Sum	200	54		1
missing	16	0		

C